4-15-2019

The Role of Statistics in Forensic Science

Karen Kafader
University of Virginia

The Role of Statistics in Forensic Science

Disciplines

Forensic Science and Technology

Comments

Posted with permission of CSAFE.

The Role of Statistics in Forensic Science

Karen Kafadar

Department of Statistics

University of Virginia

kkafadar@virginia.edu

Purdue University

15 April 2019

Acknowledgements: CSAFE, LJAF

OUTLINE

1. Statistical methods in science
2. A statistical success in Forensic Science:
   Interpreting DNA evidence (NRC 1996)
3. Statistics in Forensic Science post-facto
   CABL (NRC 2004), Anthrax (NRC 2009), EWI (NRC 2014)
4. Where statistics *can be used* in Forensic Science
   Trace & Pattern evidence, EWI experiments, Interpretation
5. From Problems to Research to Solutions to Implementation
6. Final comments: Broad Role of Statistician

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

1. Statistical Methods in Science

Science of analyzing data, characterizing uncertainties

- **Biology**: extinction/abundance of species; characterizing genetic expression (millions of SNPs) in response to stimuli; associating genotypes with phenotypes
- **Chemistry**: discovery of argon (Lord Rayleigh); source attribution via MSMS (mass spec); environmental contamination levels (San Juan River contamination)
- **Physics**: data analysis of high-energy physics (HEP) experiments to discover new particles; estimating 'big G' with uncertainty; global warming (IPCC)
- **Medicine**: clinical trials of new drugs; evaluation of treatment and screening programs; estimating disease prevalence, incidence, spread

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

Statistics in Forensic Science is notably absent from this list.

- To date, the Innocence Project has exonerated 356 people in U.S. by DNA testing (`http://www.innocenceproject.org`)
- Mistaken eyewitness identification has contributed to over 70% of these exonerations.
- "*Misapplication of forensic science is the second most common contributing factor to wrongful convictions, found in nearly half (45%) of DNA exoneration cases.*"

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

Statistics in Forensic Science:
A success: Interpreting DNA evidence

- "DNA-1" (NRC 1992) lacked statistical credibility
- "DNA-2" (NRC 1996): Statisticians' participation
- Marker selection: sensitivity (how well alleles make correct id), specificity (how well alleles distinguish individuals)
- 13 core loci (L_j, $j = 1, ...13$), each with 6–21 alleles (k_j alleles, frequency $> 0.01 \Rightarrow n_j \approx k_j(k_j + 1)/2$ genotypes at each locus)
- Calculate probabilities of "match" at 13 (now 20) loci if samples come from different sources
- "Independence": Assume outcome (genotype ID) at locus i is *independent* of outcome at locus j

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

	CSF1PO	FGA	TH01	TPOX	vWA
#alleles	8	21	6	7	9
#genotypes	36	231	21	28	45

	D3S1358	D5S818	D7S820	D8S1179
#alleles	8	8	8	10
#genotypes	36	36	36	55

	D13S317	D16S539	D18S51	D21S11
#alleles	7	7	15	17
#genotypes	28	28	120	153

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

Why DNA analysis is a successful forensic method:

- Well-defined markers (not just any 13 loci)
- Well-characterized error rates:
  High sensitivity: P{'match' | samples from same source}
  High specificity: P{'no match' | different sources}
- $\Rightarrow$ High Positive/Negative Predictive Value:
  PPV = P{samples came from same source | 'match' call}
  NPV = P{samples came from different sources | 'no match'}
- Well-designed experiments & careful analysis of experimental data to validate performance
- Well-defined procedures for execution
- Clear guidelines for interpreting/reporting results

Statistics involved in all steps; challenges remain

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

- **Statisticians working with geneticists**
  (D.P. Byar: "*A statistician working alone is a statistician making mistakes*")
- **We do *not* expect that *all* forensic methods will have the same high accuracy as DNA**
- **We *do* expect that statistics can *contribute* to *characterizing* sources of uncertainty in the methods and begin to *quantify* their effects on accuracy**
- **Statisticians must work with Forensic Scientists**
- **Goal of Scientific Method: Continuously update knowledge**

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

Where Statistics might have been used in Forensic Science: CBLA

- Scenario: Crime → evidence → bullets
- Gun recovered: match striations on bullet and gun barrell
- *No gun*: **Comparative Bullet Lead Analysis** (CBLA)
- "Working hypothesis": chemical concentration of lead used to make "batch" of bullets provides "unique signature" ⇒ "equal" concentrations of elements in Crime Scene (CS) bullets and Potential Suspect (PS) bullets may indicate "guilt"
- Local police dept sends CS, PS bullets to FBI lab
- FBI measures (in triplicate) concentrations of 7 elements

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

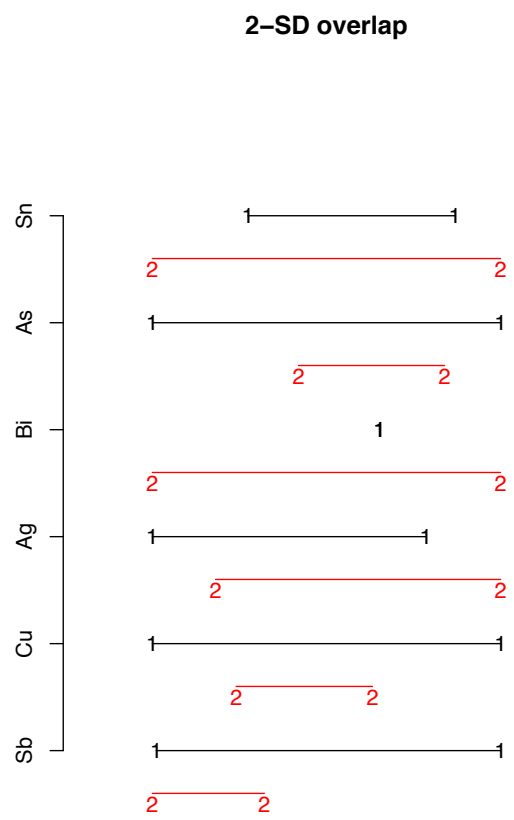
- Reports "analytically indistinguishable concentrations" between CS and PS bullets if "mean ± 2·SD intervals overlap for *all* 7 elements" (**2-SD-overlap**), provides court testimony when requested (As, Sb, Sn, Bi, Cu, Ag, Cd)
- FBI "validates" process on "1837-bullet database": "*one specimen from each combination of bullet caliber, style, and nominal alloy class was* **selected**" for database; found 693 "matches" out of (1837·1836/2) = 1,686,366 pairs of bullets
- i.e., **bullets selected to be different (not representive)**, so actual false probability rate is higher than 0.04%

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

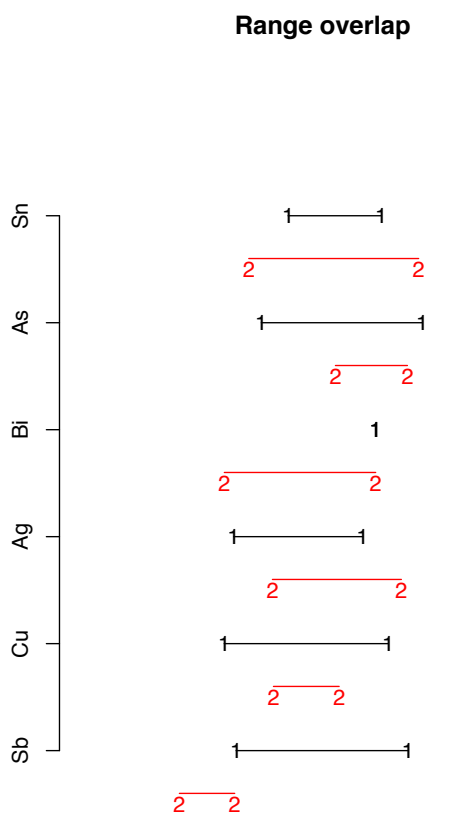
Federal bullet F001

	icpSb	icpCu	icpAg	icpBi	icpAs	icpSn
a	29276	285	64	16	1415	1842
b	29506	275	74	16	1480	1838
c	29000	283	66	16	1404	1790
mean	29260.7	281.0	68.0	16	1433.0	1823.3
SD	253.4	5.3	5.3	0	41.1	28.9
mean-2SD	28754.0	270.4	57.4	16	1350.8	1765.5
mean+2SD	29767.4	291.6	78.6	16	1515.2	1881.2
minimum	29000	275	64	16	1404	1790
maximum	29506	285	74	16	1480	1842

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org



‘Analytically indistinguishable’



All elements analytically indistinguishable except Sb

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

Statisticians on NRC Committee (report, 2004):

- "hypothesis" *bullets came from same box*:
  not sensible (manufacturing process: bullets from different batches in same box, bullets from same batch in many boxes)
- "hypothesis" *bullets came from batch with similar signature*:
  feasible (2-sample test on means) — but is it probative?
- FBI's "error rate": **selected** 1837 bullets from "70,000-bullet database" (17,000?) **to be as different as possible**
- FBI found only 693 "matches" out of $(1837 \cdot 1836/2) = 1{,}686{,}366$ pairs of bullets (0.04%)
- Simulation demonstrated otherwise: Suppose difference in concentrations in each of 7 elements is δ; what is the probability of the 2-SD test claiming a match?

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

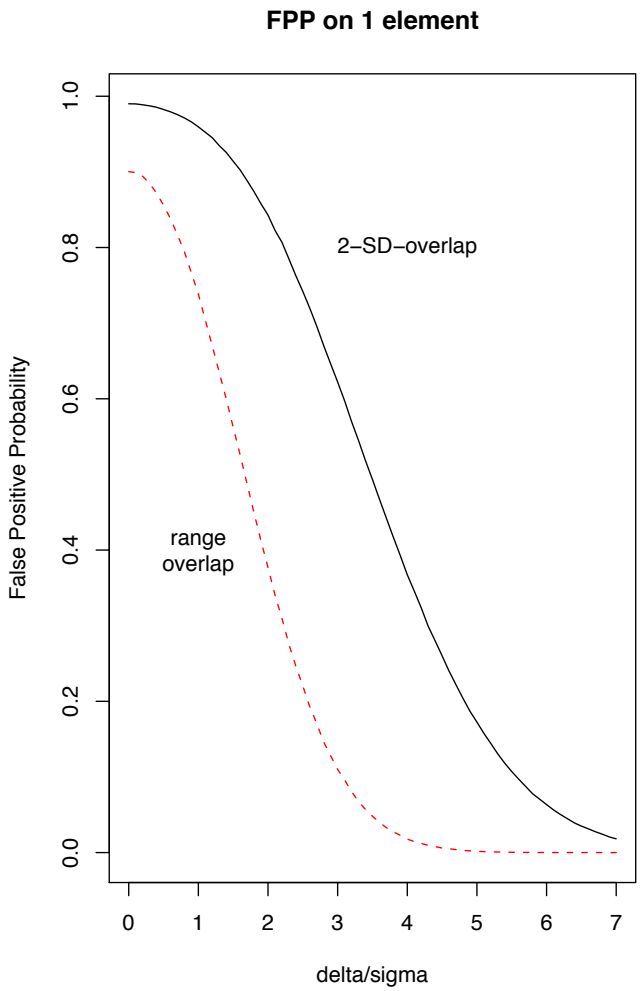
- "Innocent until proven guilty" $\Rightarrow$
  H_0: $|\mu_{CS} - \mu_{PS}| > \delta_0$, H_1: $|\mu_{CS} - \mu_{PS}| \leq \delta_0$
- Proper test: Equivalence Hotelling's T^2, not "2-SD overlap"
- Historical data $\Rightarrow$ *correlated* measurement errors
- Simulations $\Rightarrow$ "2-SD-overlap" false positive rate $> 0.04\%$

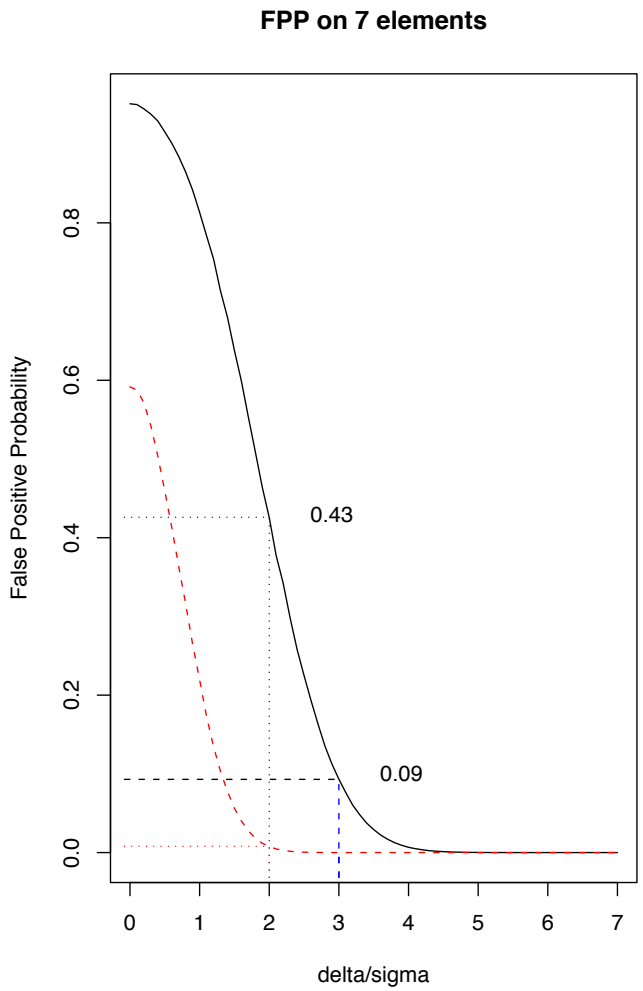
Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

Estimated correlation matrix (200 Federal bullets)

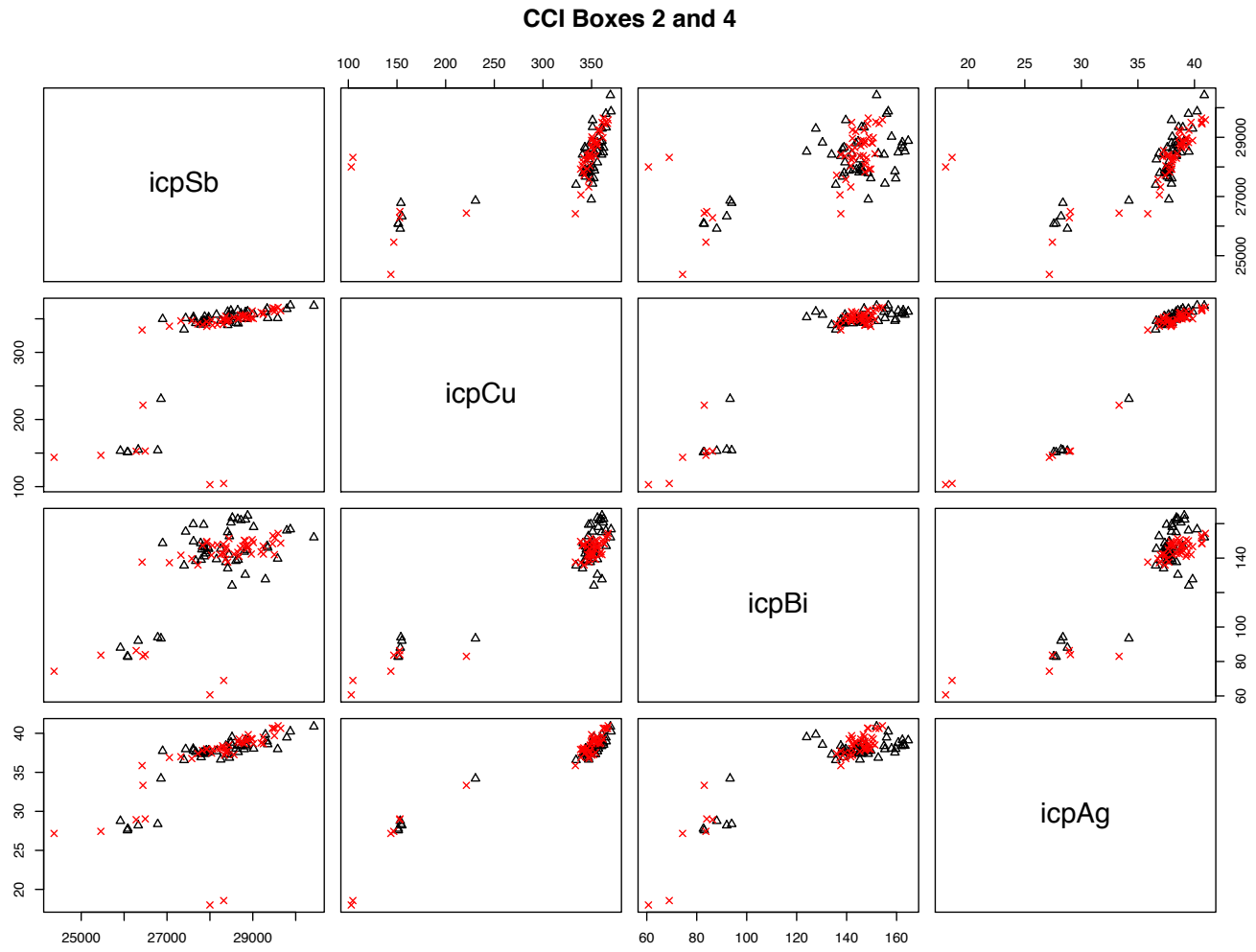
	As	Sb	Sn	Bi	Cu	Ag	(Cd)
As	1.000	0.320	0.222	0.236	0.420	0.215	0.000
Sb	0.320	1.000	0.390	0.304	0.635	0.242	0.000
Sn	0.222	0.390	1.000	0.163	0.440	0.154	0.000
Bi	0.236	0.304	0.163	1.000	0.240	0.179	0.000
Cu	0.420	0.635	0.440	0.240	1.000	0.251	0.000
Ag	0.215	0.242	0.154	0.179	0.251	1.000	0.000
(Cd)	0.000	0.000	0.000	0.000	0.000	0.000	1.000

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org





Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org



Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

For CCI boxes 2 and 4:

(∗) Prob{bullets come from different boxes | FBI 'match' }

Box 2: 674 'matches' from $(50)(49)/2 = 1225$ comparisons

Box 4: 573 'matches' from $(50)(49)/2 = 1225$ comparisons

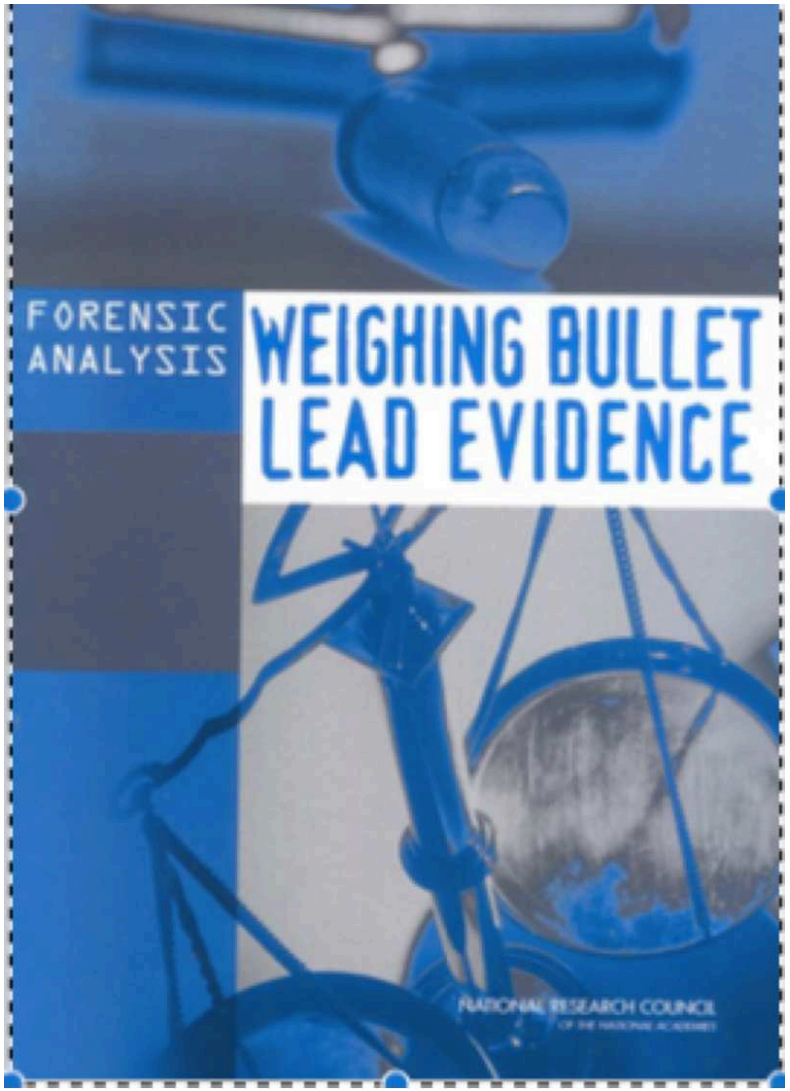
Boxes 2 & 4: 1092 'matches' from 2500 comparisons

$$(*) = \frac{P\{match|\text{dif box}\} \cdot P\{\text{dif box}\}}{P\{match|\text{dif}\} \cdot P\{\text{dif}\} + P\{match|\text{same}\} \cdot P\{\text{same}\}}$$

$$\frac{(1092/2500) \cdot (2500/4950)}{(1092/2500) \cdot (2500/4950) + (674+573)/(1225+1225) \cdot (2450/4950)}$$

$= 0.4668 \Rightarrow$ "Match" does not mean "same box"!

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org



Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

- NRC report released February 2004
- March 2005: NJ appeals court overturns 1997 murder conviction based on NRC report "raised new questions about the technique the FBI has used for decades to match bullets to crimes" (*Assoc Press*, 8 Mar 2005, p.A08)
- Sept 2005: FBI abandons CABL
  "The FBI said its decision to drop the tests was significantly influenced by the fact that 'neither scientists nor bullet manufacturers are able to definitively attest to the significance of an association made between bullets in the course of a bullet lead examination' " (*AP*, 1 Sep 2005)

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

More to come: 3 ASTM Standards for Forensic Glass

1. **XRF**: ASTM E2926-13, *Standard Test Method for Forensic Comparison of Glass Using Micro X-ray Fluorescence (μ-XRF) Spectrometry* (approved)
2. **ICP-MS**: ASTM E2330-12, *Standard Test Method for Determination of Concentrations of Elements in Glass Samples Using Inductively Coupled Plasma Mass Spectrometry (ICP-MS) for Forensic Comparisons*
3. **LA-ICP-MS**: ASTM E2927-16, *Standard Test Method for Determination of Trace Elements in Soda-Lime Glass Samples Using Laser Ablation Inductively Coupled Plasma Mass Spectrometry for Forensic Comparisons* (approved)

12–17 elements recommended in each standard

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

ICP-MS, §10: "**Calculation and Interpretation of Results**"

1. *For the Known source fragments, using a minimum of 3 measurements, calculate the mean for each element.*
2. *Calculate the standard deviation for each element. This is the Measured SD.*
3. *Calculate a value equal to 3% of the mean for each element. This is the Minimum SD.*
4. *Calculate a match interval for each element with a lower limit equal to the mean minus 4 times the SD (Measured or Minimum, whichever is greater) and an upper limit equal to the mean plus 4 times the SD (Measured or Minimum, whichever is greater).*

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

5. *For each Recovered fragment, using a minimum of 3 measurements, calculate the mean concentration for each element.*

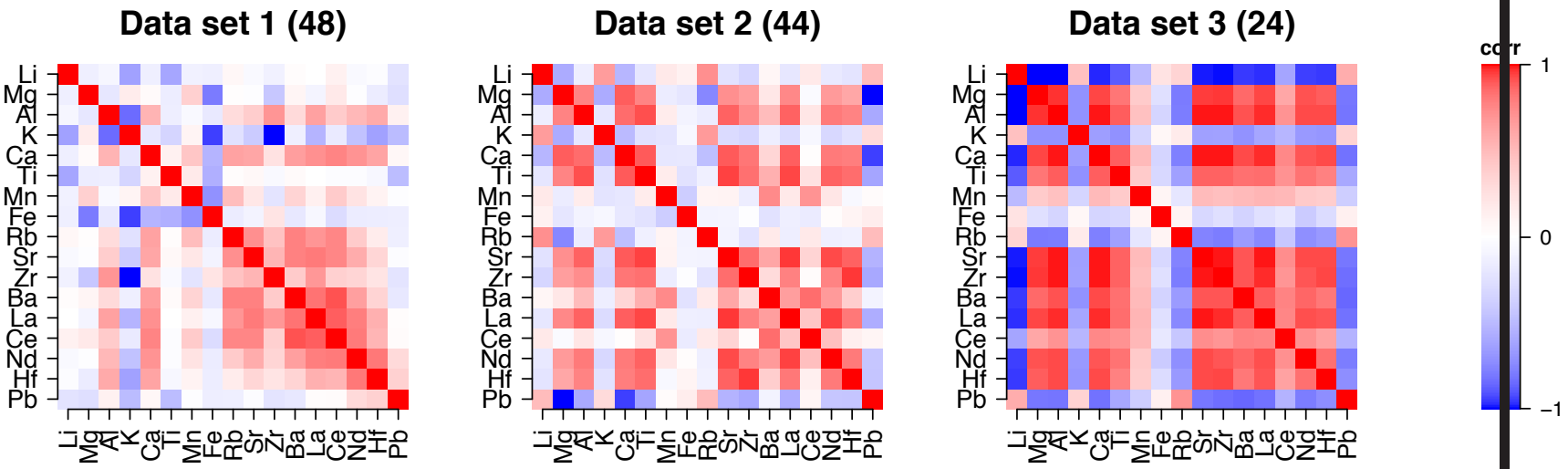
6. *For each element, compare the mean concentration in the Recovered fragment to the match interval for the corresponding element from the Known fragments.*

7. *If the mean concentration of one (or more) element(s) in the Recovered fragment falls outside the match interval for the corresponding element in the Known fragments, the element(s) does not "match" and the glass samples are considered distinguishable."*

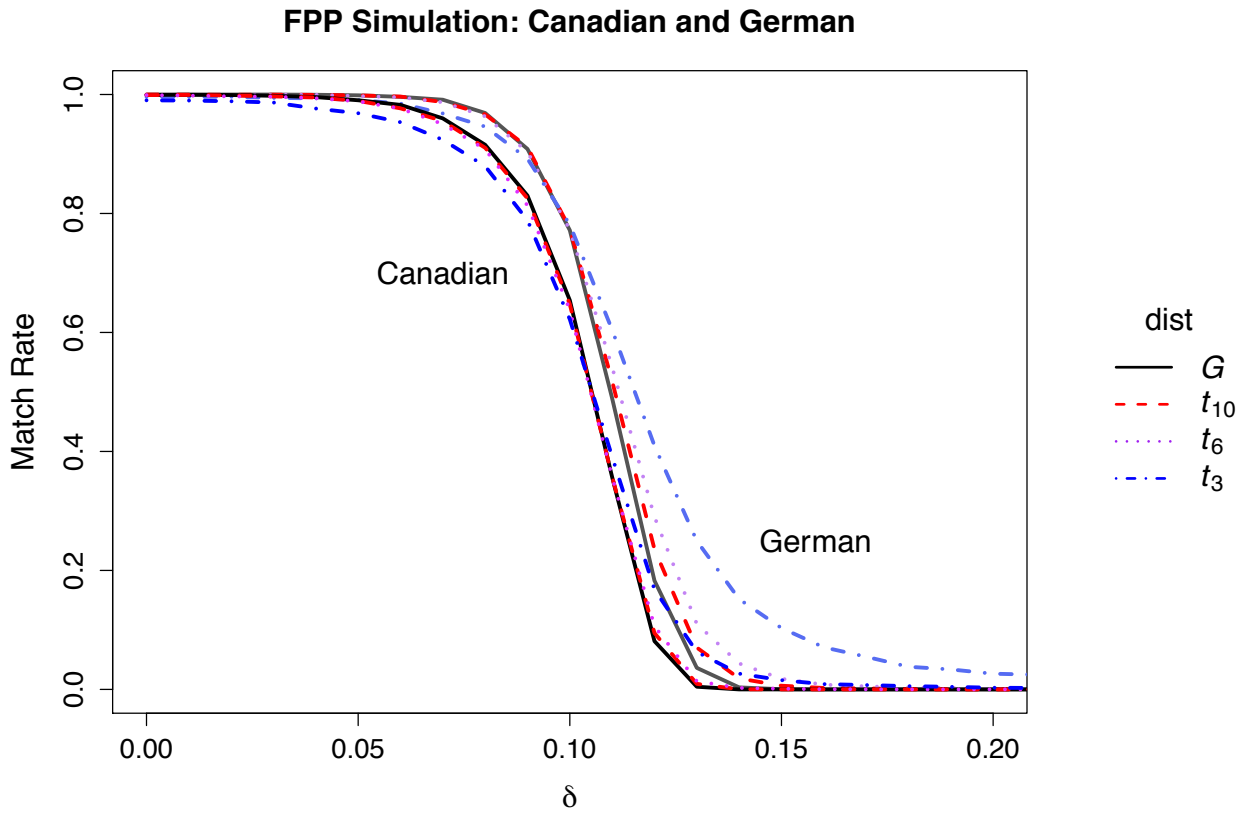
Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

- XRF, ICP-MS, LA-ICP-MS: yield good measurements
- Many sources of variability: measurement σ_r, between fragments (same pane) σ_f, between panes (close in time) σ_t, between panes (different times/manufacturer) σ_p
- Most studies: 3 reps (Germany: 6; Canada: 9)
- Very few data sets measure several fragments from *same* pane several times, across several days
- Correlated elements: Not 17 independent features (ex: Hf & Zr in same place in Periodic Table)
- Correlation matrix seems to depend on Lab

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org



Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org



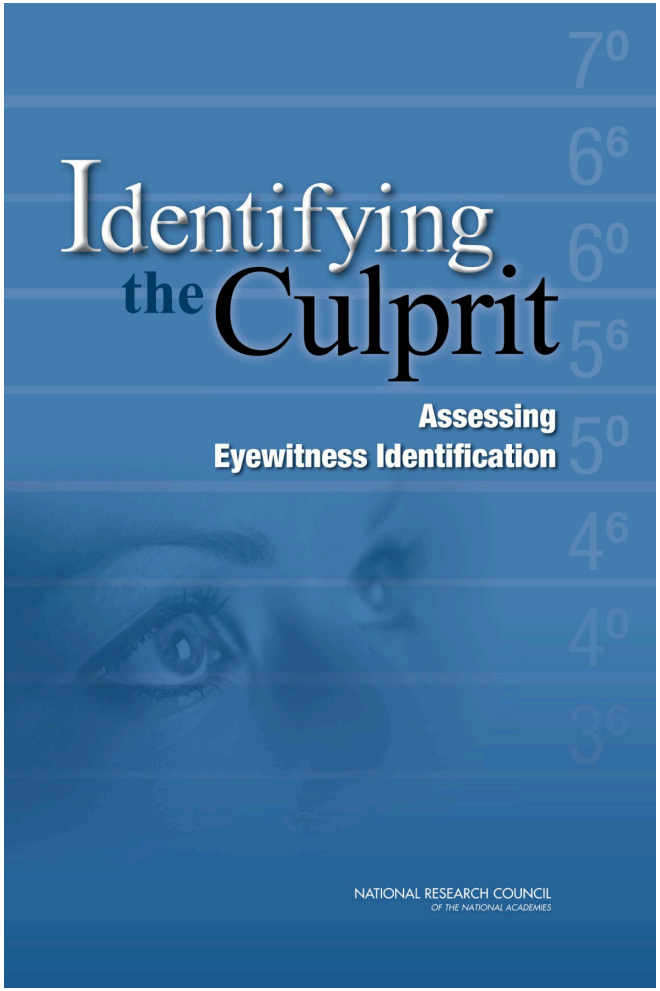
Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

Eyewitness Identification

Background:

- Eyewitness testimony can be very useful and incredibly powerful in the courtroom
- But memory is not always accurate nor reliable
- Innocence Project: 356 exonerations since 1989 from post-conviction DNA testing; 240 (72%) involved mistaken eyewitness identification (`http://innocenceproject.org`)
- What procedures are used in eyewitness identification (EWI)?
- Which procedures lead to accurate identifications?
- **How to compare procedures in terms of accuracy?**

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org



Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

The task: Identify person in the incident (assault, robbery, ...)

Binary decision, binary outcome

		Witness Classification	
		"Guilty"	"Innocent"
True Status	Guilty	True +	False –
of Suspect	Innocent	False +	True –

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

Ronald Cotton & Jennifer Thompson: *Picking Cotton*

- 1984 rape of Jennifer Thompson (college student in NC)
- Police sketch → Ronald Cotton
- 6 photos; Jennifer reluctantly chooses 2, then 1:
  "I think this is the guy."
- Detective: "You're sure?" — "Positive. Did I do OK?"
- Live lineup: Only Cotton was repeated from photo lineup
- Thompson selects Cotton: "looks the most like him"
- Courtroom: "100% sure. That's the guy who raped me."
- Convicted to life in prison + 54 years
- 1995: Cotton exonerated by DNA; Police arrest Bobby Poole.

NAS report, p10

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

John Jerome White

- Victim states: Attacker was "well built", "round face"
- 5-person live lineup: Selects White (middle)
- Courtroom: "*Do you see a person in the courtroom here today that was the person who came in your apartment that night?*"
- Victim: "*That's him (indicating).*"
- White convicted; 22+ years in prison; DNA exoneration 2007

B.L. Garrett, `http://harvardpress.typepad.com/hup_publicity/2011/03/understanding-eyewitness-misidentifications.html`

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

Situational aspects of EWI (*Estimator variables*):
Beyond the control of the criminal justice system

1. Eyewitness' level of stress or trauma at incident
2. Conditions affecting visibility
3. Distance between witness and perpetrator
4. Presence/absence of threat (e.g., weapon)

etc.

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

Procedural aspects of EWI (*System variables*):

1. Conditions & protocols for **lineups**
   (e.g., *sequential vs simultaneous; fair vs biased*)
2. Nature of instructions (oral or written, short or long, ...)
3. Presence/absence of feedback
4. Number and similarities of fillers with "target"
5. Retention interval (longer ⇒ less reliable)

etc. *Which factors matter most to accuracy?*

Focus: Compare accuracy between two lineup procedures —
but methods should apply to comparing *any* two procedures

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

"That's him—the one on the right."

From THE NEW YORKER, March 7, 2011

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

Sequential vs Simultaneous?

- *Sequential*: Present each photograph, one at a time
- *Simultaneous*: Present all six photographs at once
- Early research: "Sequential is more accurate"
- Later research: "Metric for comparison is incomplete; Simulaneous is more accurate"
- Which was correct?

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

Lab tests and proposed metrics

Lab tests: Present participants (usually Psych 1 students) a scenario, followed by lineup (sequential or simultaneous); count proportions of correct IDs (HR = *hit rate*) and mistaken IDs (FAR = *false alarm rate*)

1. *Diagnosticity Ratio*: Collapse all participants, all scenarios:

$$\begin{aligned} &\textit{diagnosticity ratio} = \textit{hit rate / false alarm rate} \\ &= \textit{Sensitivity / (1 – Specificity)} \ (= LR^{+}) \end{aligned}$$

2. Some participants express more *confidence* in their choices; *confidence* is related to *accuracy*; therefore, we should look at HR and FAR as functions of *levels of expressed confidence*.

Which approach is correct?

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

Probably neither. Use logistic regression! Or another binary classifier: Outcome = Right or Wrong; Covariates lineup type, presence/absence of weapon, good/poor lighting, ...

- *Sensitivity*: When shown the *true* perpetrator, what is the probability that the "witness" identifies him/her?
- *Specificity*: When shown an *imposter*, what is the probability that the "witness" excludes him/her?
- *Sensitivity*, *Specificity* can be estimated only in studies *where truth is known* (by design)
- Real life: All you have is response:
  "Yes, that's the one" or "No, not that one"

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

- *Positive Predictive Value (PPV)*: If claim is "Yes, that's the one": Probability that identified person is the perpetrator?
- *Negative Predictive Value (NPV)*: If claim is "No, not the one": Probability that excluded person is **not** the perpetrator?
- *PPV, NPV* are functions of *Sensitivity, Specificity, and odds that the suspect is the true perpetrator*
- *Higher Diagnosticity Ratio* $(LR_+) \Rightarrow$ *higher* PPV:

$$PPV = 1 \,/\, (1 + Odds/DR)$$

- Correct **exclusions**? NPV is related to $LR_- = (1 - sens)/spec$: $NPV = 1 \,/\, (1 + LR_- \cdot OR)$
- Alice Liu: Better ways to combine LR_+ **and** LR_- (Dx medicine)

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

Relationship between Confidence & Accuracy?

If you believe confidence is related to accuracy:

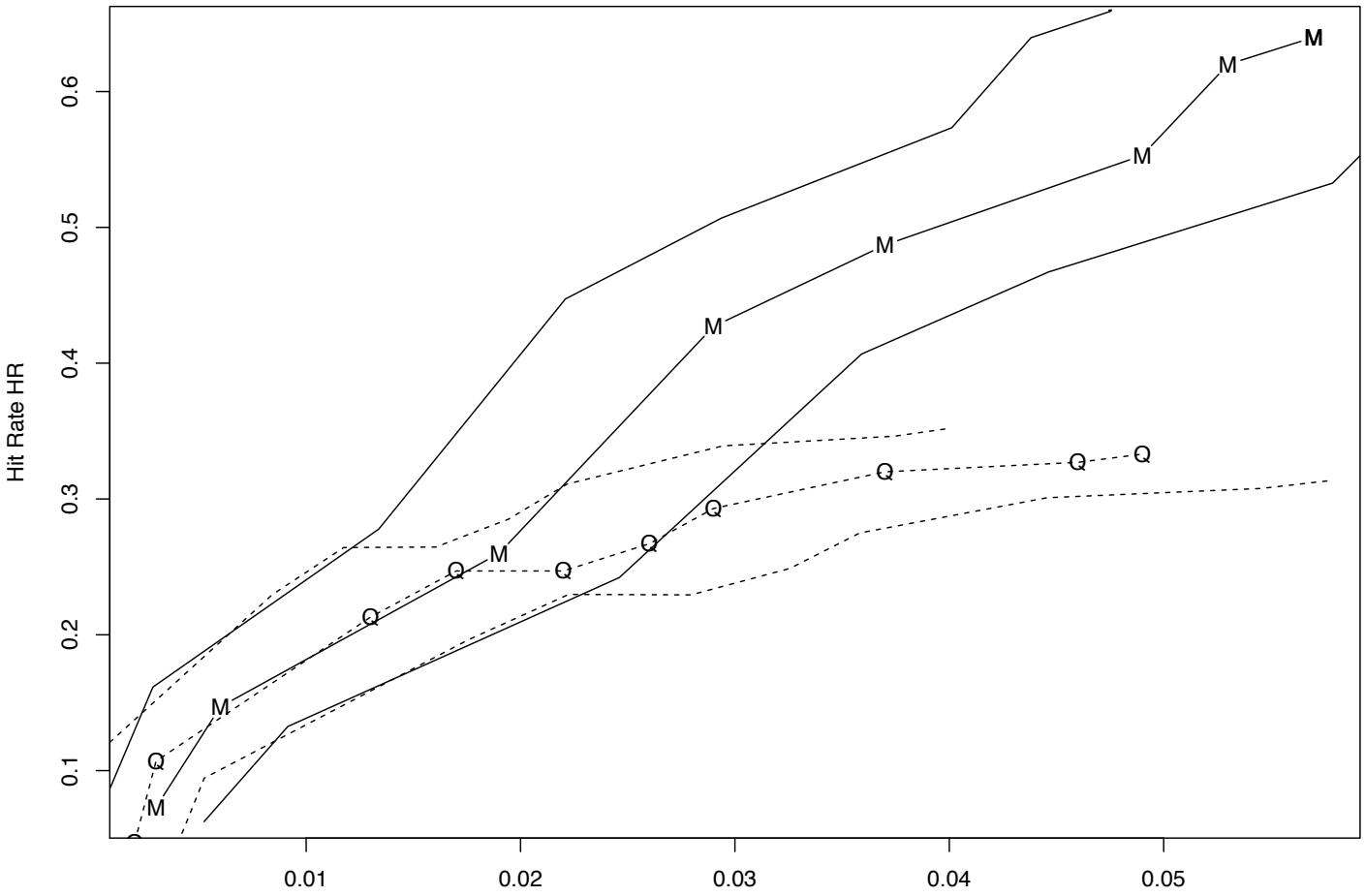
- consider calculating $DR = HR/FAR$ as a function of *Expressed Confidence Level (ECL)*
- Split the sample participants into categories of ECL (those who expressed 10%, ..., 90% confidence); calculate DR for each ECL category
- even better: Plot HR vs FAR for different ECLs
- ROC curve = Receiver Operating Characteristic
- Quality control, comparing medical diagnostic procedures

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

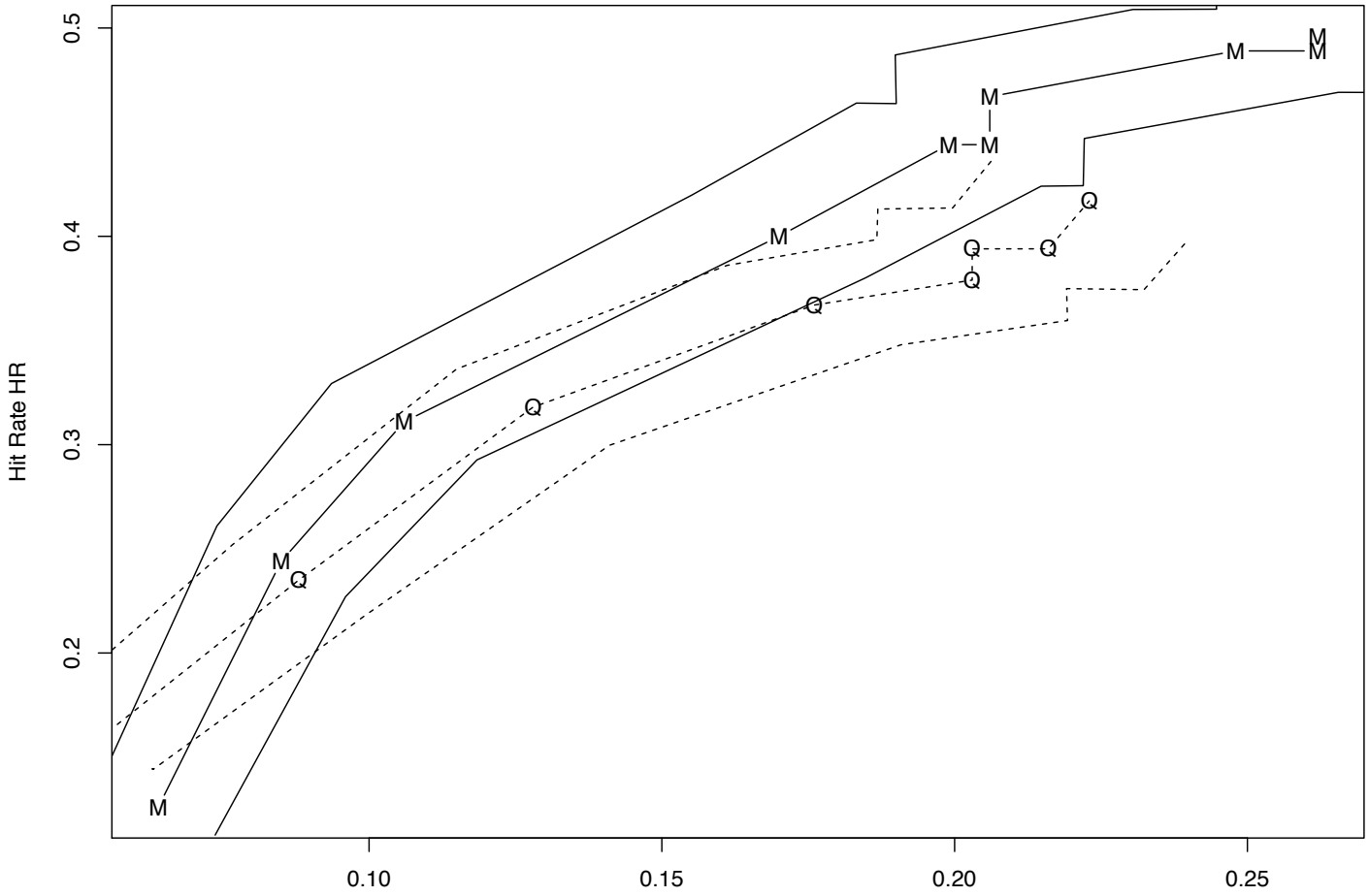
Problem: Data points (HR, FAR) have uncertainty!

- John Tukey (in discussing uncertainty in rates at NCI):
  "*What has happened is history. What might have happened is science and technology. So what you are really interested in is what might have happened if you could do it all over again.*"
- Simulate what would happen if you calculated all the HRs and FARs (for different ECLs) *as if* you repeated the same experiment all over again
- DR vs ECL for Sequential and for Simultaneous:
  How different are they?
- How different do the two ROC curves look for Sim vs Seq?
- Resulting uncertainty is underestimated, because ECLs can change (e.g., "40%" today; "20%" tomorrow)

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org



Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org



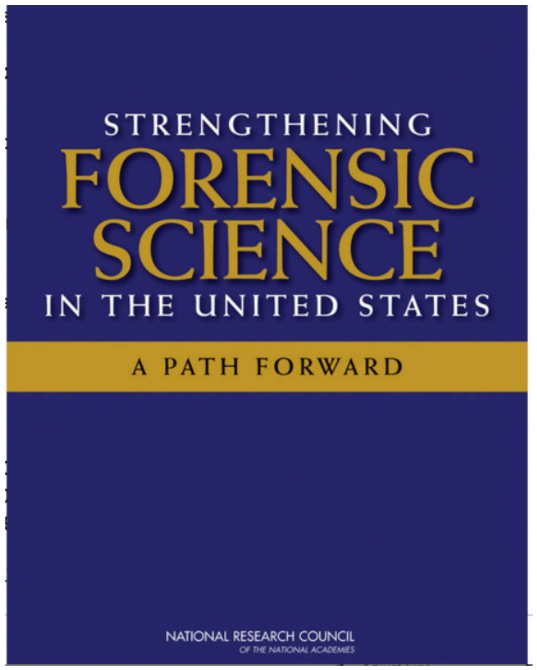
Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

The Broader Picture

2007: National Academies' Committee on Science, Technology, and Law convened committee to study reliability of Forensic Science Methods generally, *except DNA*, including:

- Trace evidence (bullet lead, glass, tape, paint, ...)
- Toxicology (drug analysis, ...)
- Pattern Evidence (fingerprints, shoe prints, tire tracks, blood pattern analysis, handwriting analysis, bite marks, ...)
- Crime scene evidence (arson, hairs & fibers, ...)
- Digital evidence

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org



"*With the exception of nuclear DNA analysis, however, no forensic method has been rigorously shown to have the capacity to consistently, and with a high degree of certainty, demonstrate a connection between evidence and a specific individual or source*" (p7)

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

2009 NAS report calls for reforming forensic science

(Co-chairs: Judge H.T. Edwards, Constantine Gatsonis)

- Many forensic disciplines lacked validation studies
- Only DNA inference based on validated probability model from which error rates can be estimated
- Claims of "zero error rate" not plausible
- Pattern evidence disciplines (latent prints, shoeprints, tire tracks, ballistics, tool marks, ...) address common issues but examiners from them do not share 'lessons learned' with others
- **Report called for more research & better coordination with scientists & statisticians**

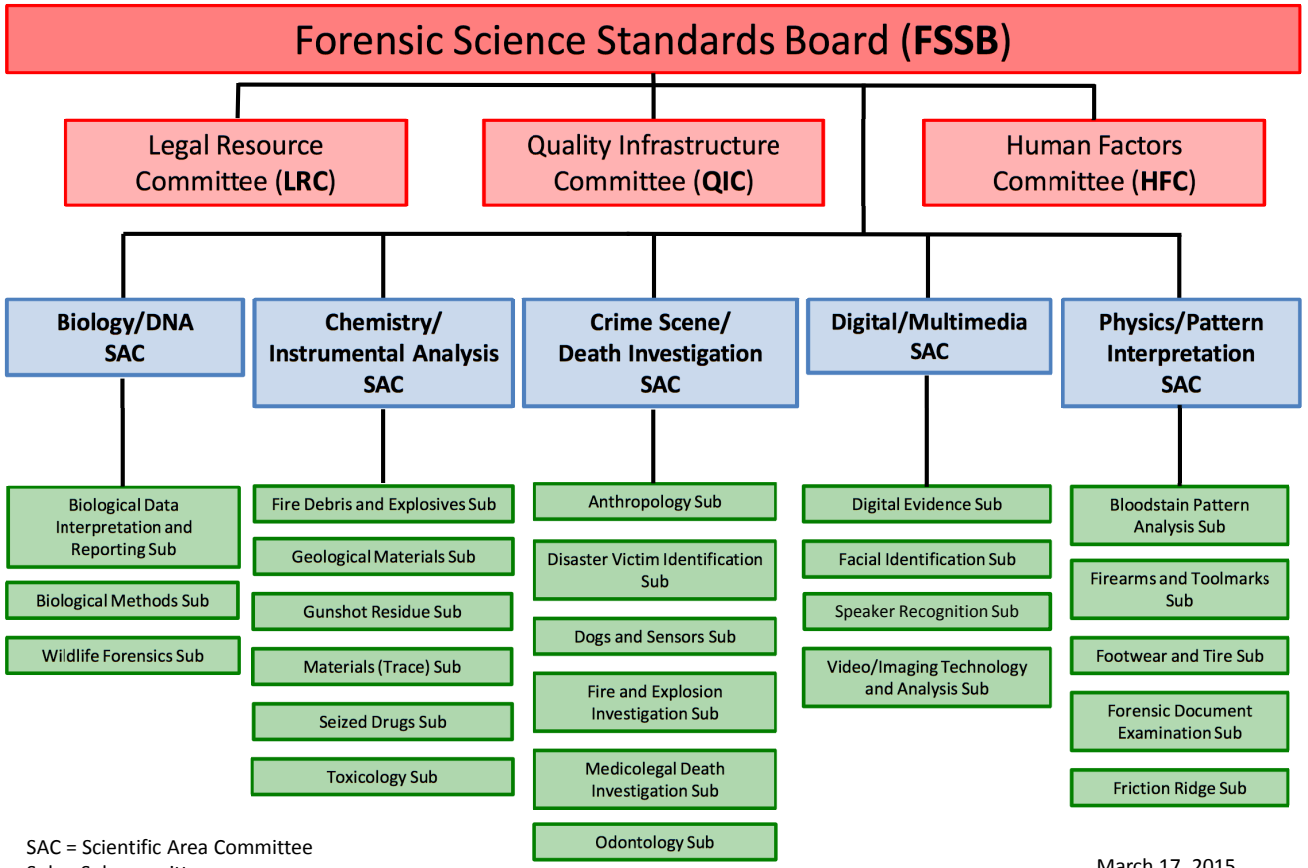
Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

Government response:

- National Commission on Forensic Science (NCFS): 13 mtgs: `https://www.justice.gov/archives/ncfs` (NIJ-NIST): Assembled research scientists, forensic practitioners, judges, legal scholars (AG disbanded in April 2017)
- Organization of Scientific Area Committees (OSAC): 25 FS disciplines organized into 5 Scientific Area Committees (NIST-NIJ): Approve guidelines and standards
- NIST RFP for Center of Excellence in Forensic Science (Cooperative agreement): Conduct research to strengthen research and ties between researchers & practitioners
- **PCAST report, Sep 2016: Endorsed NAS 2009 Report**

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

Organization of Scientific Area Committees (OSAC)



SAC = Scientific Area Committee
Sub = Subcommittee

March 17, 2015

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

Center for Statistical Applications in Forensic Evidence (CSAFE):
4 teams from university statistics departments

- Iowa State Univ: Alicia Carriquiry (Director)
- Carnegie Mellon Univ: Stephen Fienberg $\rightarrow$ Bill Eddy
- University of California-Irvine: Hal Stern
- University of Virginia: Karen Kafadar

Each team works on projects in pattern evidence deemed of importance to NIST & to forensic science: `forensicstats.org`

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

Statistics Needed in Forensic Science

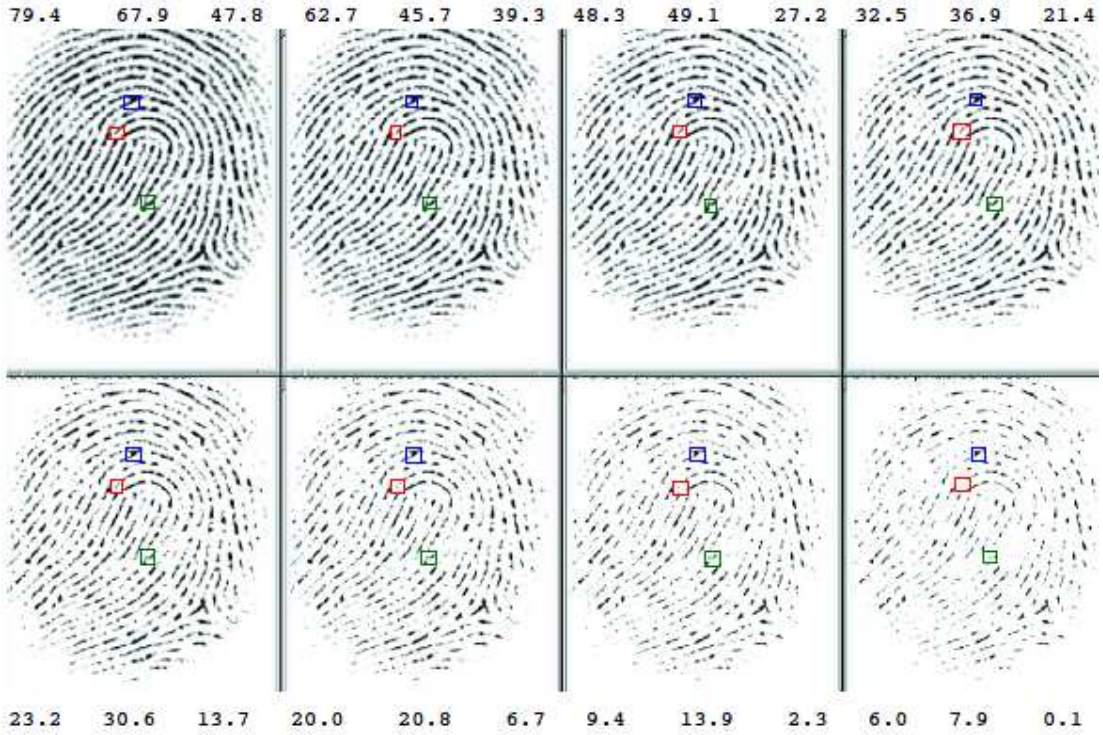
- **Problem Identification**: How to compare two items?
- **Evidence Examination**: Is the evidence suitable for examination (quality)?
- **Process Identification**: What procedures are used for comparison? Are they objective, measureable, repeatable?
- **Research**: What alternative approaches may be appropriate?
- **Comparing procedures' error rates**: How to estimate?
- **Design**: Experiments for valid comparisons of approaches?
- **Stating conclusions**: Likelihood Ratios, Posterior Odds

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

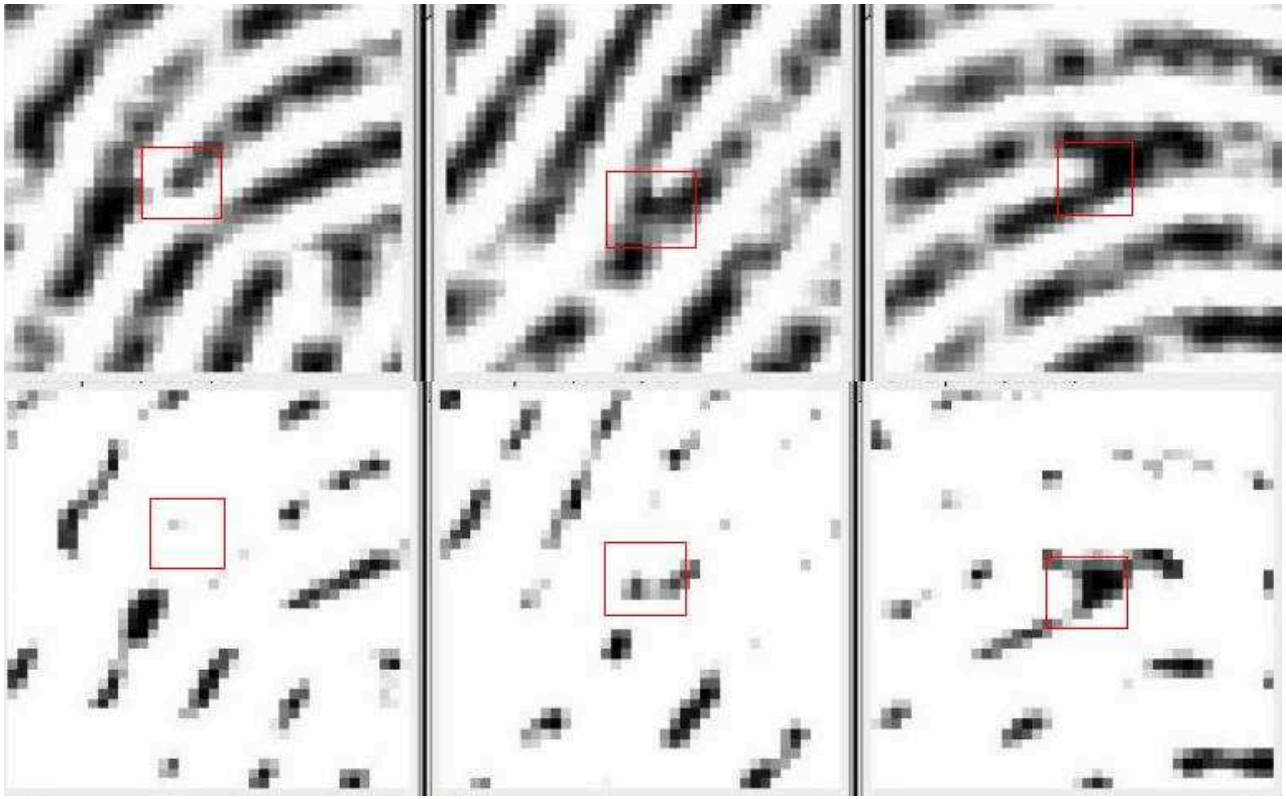
Some progress

- **FRStat**: Metric for assessing "significance" of fingerprint match (*Swofford, Liu et al. 2018, JFS*)
- Black Box Study: LPE accuracy (*Ulery et al. 2011, PNAS*)
- ISU (CSAFE): Ballistics matching
- CMU & UCI: Digital signature matching
- UVA: Quality Metrics

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org



Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org



Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

Quality scores for 3 minutiae in 8 increasingly degraded images:

Image#	left (red)	center (blue)	right (green)
1	79.4	67.9	47.8
2	62.7	45.7	39.3
3	48.3	49.1	27.2
4	32.5	36.9	21.4
5	23.2	30.6	13.7
6	20.0	20.8	6.7
7	9.4	13.9	2.3
8	6.0	7.9	0.1

(*Peskin & KK, 2017*)

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

Final Comments: Roles for Statisticians

- "Statistical thinking"
- Quantify vague concepts, sources of uncertainty
- Identify confounding factors
- Develop methods to improve accuracy in identifications
- Design validation experiments
- Interpret & communicate results of analyses

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

NAS 2009 Report Received Innocence Networks 2018 Champion of Justice Award, 4/12/2019

In accepting the award, Judge Edwards said:

"[T]o the exonerees who are here today: I am deeply pained by the indignities and personal suffering that you have endured at the hands of injustice. Most of us cannot begin to comprehend the ordeals that you have faced. It is beyond our understanding. Our system of justice failed you, and you can never get back what you lost. You have my most sincere apologies."

Judge Edwards to KK:

"So many of the exonerees have thanked me for apologizing. Many said that no public official had ever apologized before. Sad."

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

Judge Edwards: *"In his 1963 Letter from Birmingham Jail, the Rev. Martin Luther King, Jr., reminded us that '[i]njustice anywhere is a threat to justice everywhere.' We are not talking about good science merely for its own sake. We are talking about the need for good science in order to serve justice ... that will help us to avoid wrongful convictions like those suffered by the exonerees who are with us today. Goodness, commitment, resources, and intelligent effort can get it done. And when justice is done, our society as a whole is better for it."*

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

References

Dorn H, Ruddell D, Heydon A, Burton B (2015), Discrimination of float glass by LA-ICP-MS: assessment of exclusion criteria using casework samples, *Can Soc Forensic Science J* 48(2): 85-96

Kafadar K (2015), Statistical Issues in Assessing Forensic Evidence, *Int Statistical Rev* 83(1), 111-134.

Spiegelman CS, Kafadar K: The Case of Bullet Lead Data as Forensic Evidence, *Chance* 19(2):17–25 (2006)

Pan KDH; Kafadar K (2018), Statistical modeling and analysis of trace element concentrations in forensic glass evidence, *Annals of Applied Statistics*

Weis P, Dückling M, Watzke P, Menges S, Becker S (2011), Establishing a match criterion in forensic comparison analysis of float glass using laser ablation inductively coupled plasma mass spectrometry, *J Anal Atomic Spectroscopy* 26:1273-1284.

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org